A SHORT PROOF OF THE BOUNDED GEODESIC IMAGE THEOREM

RICHARD C. H. WEBB

Abstract. We give a combinatorial proof, using the hyperbolicity of the curve graphs, of the bounded geodesic image theorem of Masur–Minsky. Recently it has been shown that curve graphs are uniformly hyperbolic, thus a universal bound can be given for the diameter of the geodesic image. We also generalize the theorem for projections to markings of the whole surface.

1. Introduction

We write $S_{g,p}$ to denote the genus g surface with p points removed and $\xi(S) = 3g-3+p$ to denote the *complexity* of $S = S_{g,p}$. We say a simple closed curve on S is *essential* if it does not bound a disc or once-punctured disc. In general, we say that an isotopy class of some subset of S *misses* another isotopy class of some subset if they admit disjoint representatives, and otherwise we say that they *cut*. A *curve* is an isotopy class of essential simple closed curve. We write $\mathcal{C}(S)$ to denote the *curve graph* of S, whose vertex set is the set of curves on S with edges between non-equal curves that miss; this is the 1-skeleton of the *curve complex* which was introduced by Harvey [6]. Throughout, $S = S_{g,p}$ with $\xi(S) \geq 2$. For the surfaces $S_{0,4}$ and $S_{1,2}$, one can use the Farey graph, the description of its geodesics and a lifting argument to prove Theorem 3.2.

We shall abuse notation by simply writing γ to mean both the simple closed curve γ and its isotopy class. We write d_S to denote the path metric on $\mathcal{C}(S)$ with unit length edges. A sequence of curves $g = (\gamma_i)$ is a *geodesic* if for all $i \neq j$, we have $d_S(\gamma_i, \gamma_j) = |i-j|$. We say $\mathcal{C}(S)$ is δ-hyperbolic if for all geodesic triangles g_1, g_2, g_3, we have $g_1 \subset N_\delta(g_2 \cup g_3)$, where N_δ is the metric closed δ-neighbourhood.

Theorem 1.1 ([12])**.** *Fix $S = S_{g,p}$ with $\xi(S) \geq 2$. There exists $\delta \geq 0$ such that $\mathcal{C}(S)$ is δ-hyperbolic.*

We write *subsurface* to denote a compact, connected, proper subsurface of S such that each component of its boundary is essential in S. Throughout, we do not consider subsurfaces that are homotopy equivalent to $S_{0,3}$; in this case Theorem 3.2 is straightforward.

For a non-annular subsurface $Y \subset S$. We write ∂Y for the boundary of Y. We now define a map $\pi_Y : \mathcal{C}_0(S) \to \mathcal{P}(\mathcal{AC}_0(Y))$, where $\mathcal{AC}(Y)$ is the *arc and curve complex* of Y, and generally $\mathcal{P}(X)$ is the set of subsets of X. Given a curve $\gamma \in \mathcal{C}(S)$, isotope γ so that it intersects Y minimally. We define $\pi_Y(\gamma)$ to be the arcs and/or curves $\gamma \cap Y \subset Y$. This is non-empty if and only if γ cuts Y. The map π_Y is the *subsurface projection* to the arc and curve complex of Y. We write $\pi_Y(A) = \cup_{\gamma \in A} \pi_Y(\gamma)$.

When Y is an annulus we write ∂Y for the core curve of Y. This core curve represents a subgroup of $\pi_1(S)$ and therefore there is an associated cover $p_Y : S_Y \to S$, where S_Y is homeomorphic to the interior of an annulus. There is a homeomorphic lift of Y to S_Y which we write Y'. One can compactify S_Y to a closed annulus by using a hyperbolic metric on S. Let $\mathcal{AC}_0(Y)$ be the set of arcs that connect one boundary component of S_Y to the other, modulo isotopies that fix the endpoints. Two arcs are adjacent if they admit disjoint representatives. We write $\mathcal{AC}(Y)$ to denote this graph. Given a curve γ that cuts Y, we define $\pi_Y(\gamma)$ to be the set of arcs of the preimage $\tilde{\gamma} = p_Y^{-1}\gamma$ that connect the two boundary components of S_Y. Otherwise, $\pi_Y(\gamma) = \emptyset$. This defines the subsurface projection $\pi_Y : \mathcal{C}_0(S) \to \mathcal{P}(\mathcal{AC}_0(Y))$ when Y is an annulus.

We write $d_{\mathcal{AC}(Y)}$ to denote the standard metric on the graph $\mathcal{AC}(Y)$. We write $d_Y(A) = \mathrm{diam}_{\mathcal{AC}(Y)}(\pi_Y A)$ and $d_Y(A, B) = \mathrm{diam}_{\mathcal{AC}(Y)}(\pi_Y(A) \cup \pi_Y(B))$. The following lemma is immediate, see also [9, Lemma 2.2].

Lemma 1.2. *Let Y be a subsurface of S and let γ_1, γ_2 be curves on S. Suppose that γ_1 cuts Y, γ_2 cuts Y and γ_1 misses γ_2. Then $d_Y(\gamma_1, \gamma_2) \leq 1$.* □

We shall give a proof of the bounded geodesic image theorem of Masur–Minsky [9, Theorem 3.1]. We shall give a bound that depends only on δ, where $\mathcal{C}(S)$ is δ-hyperbolic.

Theorem 3.2. *Given a surface S there exists $M = M(\delta)$ such that whenever Y is a subsurface and $g = (\gamma_i)$ is a geodesic such that γ_i cuts Y for all i, then $d_Y(g) \leq M$.*

Recently, it has been shown that there exists δ such that $\mathcal{C}(S)$ is δ-hyperbolic for all surfaces S in Theorem 1.1, see Aougab [1], Bowditch [2], Clay–Rafi–Schleimer [4] and Hensel–Przytycki–Webb [7].

Corollary 1.3. *There exists M independent of the surface S in Theorem 3.2.*

In the last section, we describe markings on S in terms of graphs embedded in S that fill. Given a multicurve α, and a curve γ that fills with α, one can define such a graph $\Gamma_\alpha(\gamma)$. This gives a projection Γ_α to a set of markings. Our proof of Theorem 3.2 generalizes to these projections.

Theorem 4.1. *Suppose a multicurve α and a geodesic $g = (\gamma_i)$ satisfy γ_i, α fill S for each i. Then $\mathrm{diam}_{\mathcal{M}_{k,l}(S)}(\Gamma_\alpha(g)) \leq M$, where M depends on S.*

2. Loops and surgery

2.1. **Loops.** Throughout this section, α and β are both collections of pairwise disjoint, essential, simple closed curves on S such that α and β intersect minimally (equivalent to α and β do not share a bigon, see for example [5, Proposition 1.7]) and α, β fill S.

We say a collection of simple closed curves $\{\gamma_i\}$ is *sensible* if they are essential, pairwise in minimal position, and with no triple points, i.e. for distinct i, j, k, we have $\gamma_i \cap \gamma_j \cap \gamma_k = \emptyset$.

Let γ, α, β be sensible. Recall that whenever we orient γ and β arbitrarily, each point $\gamma \cap \beta$ has a sign of intersection ± 1. We say a pair of such points *have opposite sign* if the signs of intersection are non-equal, and *have same sign* otherwise. This notion does not depend on the orientation of γ, β or S.

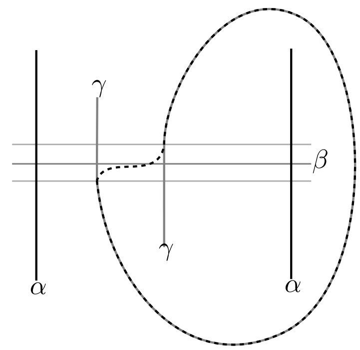

FIGURE 1. The curve γ' is dotted.

Definition 2.1. We say that γ is an (α, β)-*loop* if for each arc $b \subset \beta - \alpha$ we have $|\gamma \cap b| \leq 2$ with equality only if $\gamma \cap \beta$ have opposite sign.

Definition 2.1 is inspired by Leasure's $(\alpha \cup \beta)$-*cycles* [8, Definition 3.1.6]. These cycles allow one to construct quasigeodesics on closed surfaces with a combinatorial description. Definition 2.1 is an adaptation, which allows one to work on punctured surfaces. Both (α, β)-loops and Leasure's cycles satisfy some variant of Lemma 2.5, however cycles a priori require larger constants for Lemma 2.5 and a more careful proof since they are not necessarily in minimal position with α and β. Our surgery argument to construct (α, β)-loops from curves is necessarily more technical, but they will intersect α and β minimally.

2.2. **Surgery.** Suppose that γ, α, β are sensible. We shall describe a surgery process on γ to construct an (α, β)-loop which will be written γ'. If γ is an (α, β)-loop then we set $\gamma' = \gamma$. If γ is not an (α, β)-loop then let c be a minimal (with respect to inclusion) connected subarc $c \subset \gamma$ such that there exists an arc $b \subset \beta - \alpha$ with either

- $c \cap b$ is a pair of points with same sign
- $c \cap b$ has cardinality at least 3

Since c is minimal we have that c has endpoints on b, b is the unique arc with properties described above, and $|c \cap b| \leq 3$. Thus, each arc $b' \subset \beta - \alpha$ such that $b' \neq b$, we have $|c \cap b'| \leq 2$ with equality only if $c \cap b'$ have opposite sign.

In what follows, we write $N = N(\beta)$ to denote a closed regular neighbourhood of β. We now describe how to construct γ', in each case of how c intersects b.

Case 1: $|c \cap b| = 2$ and $c \cap b$ have same sign. See Figure 1. Write $R \subset N - \alpha$ to denote the rectangle with $b \subset R$. Let $\{p_1, p_2\} = c \cap \partial R$. Connect p_1 to p_2 by an arc $a \subset R$ that intersects b once and intersects c only at the endpoints of a. We let γ' be the simple closed curve $a \cup (c - R)$.

Case 2: $|c \cap b| = 3$ with alternating signs of intersection with respect to some order on b. See Figure 2. Let p_1, p_2, p_3 be the points $c \cap b$ in some order along b. Let $c_1, c_2 \subset c$ be arcs such that $c_1 \cup c_2 = c$, $\partial c_1 = \{p_1, p_2\}$ and $\partial c_2 = \{p_2, p_3\}$. Connect $c_1 \cap \partial R$ to $c_2 \cap \partial R$ by two disjoint arcs $a_1, a_2 \subset R$ so that a_1 intersects c_1, c_2 only at its endpoints and intersects b once, and similarly a_2. We let $\gamma' = a_1 \cup (c_1 - R) \cup a_2 \cup (c_2 - R)$.

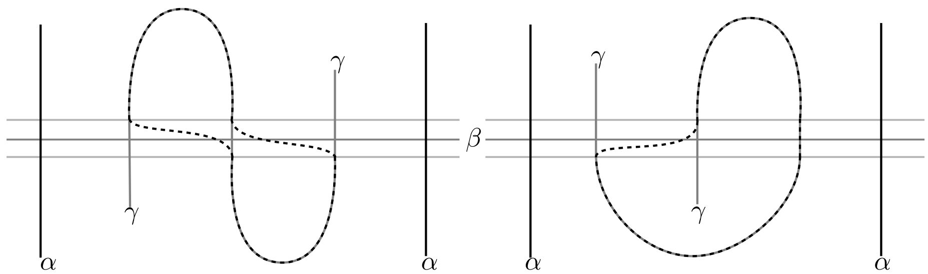

FIGURE 2. Surgery in Case 2 on the left, and Case 3 on the right.

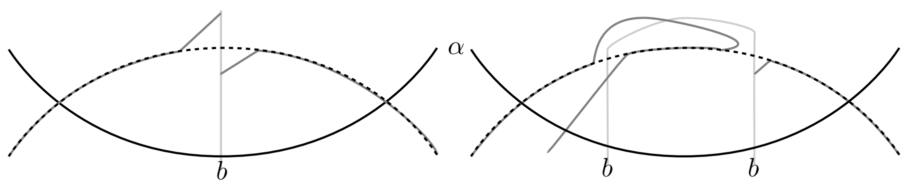

FIGURE 3. The argument in Lemma 2.2

Case 3: $|c \cap b| = 3$ with non-alternating signs of intersection. See Figure 2. We define γ' in a similar fashion as Case 1.

Lemma 2.2. *In each case above, γ' is in minimal position with β and with α. Furthermore, γ' is essential and an (α, β)-loop.*

Proof. Any arc of $\gamma' - \beta$ is isotopic in $S - \beta$ to some arc of $\gamma - \beta$. Therefore, γ' and β cannot share a bigon since γ and β do not, thus γ' is essential. We now show that γ' and α do not share a bigon in all the cases of the surgery process described above, via contradiction.

Case 1: See Figure 3. Pick an innermost bigon B between the pair γ', α. We must have the arc $a \subset \partial B$, otherwise γ and α share a bigon. We have $a \cap b \neq \emptyset$, and since γ' and β do not share a bigon, we must have one endpoint of b in ∂B. Let $\{p\} = B \cap c \cap b$, and $c_\alpha \subset \gamma - \alpha$ be the arc with $p \in c_\alpha$. Now γ and β do not share a bigon, and neither does γ intersect itself, thus c_α is contained with the disc $B \cup T$, where T is the triangle region adjacent to B cobounded by γ', γ and β. We conclude that c_α and α cobound a bigon, contradicting γ and α do not share a bigon.

Case 2: It suffices to show each connected component of $S - (\gamma' \cup \alpha)$ adjacent to at least one of the arcs a_1, a_2 is not a bigon. We start with the component containing p_2: if this is a bigon B, then the arc $b' \subset b - \gamma'$ with $p_2 \in b'$ satisfies $b' \subset B$ hence b' cobounds a bigon with γ'. This contradicts γ' and β do not share a bigon. Now we argue that the component containing p_1 is not a bigon (and similarly p_3). Suppose this component was a bigon B, first suppose that $a_1 \subset \partial B$ but $a_2 \cap \partial B = \emptyset$, then one can follow a similar argument as in Case 1. If $a_1, a_2 \subset \partial B$, then see Figure 3 on the right. A similar argument again can be given as in Case 1.

Case 3: One can argue similarly to that of Case 1. $\square$

Our Lemma 2.3 is the generalization of [8, Proposition 3.1.7].

Lemma 2.3. *Suppose $\gamma_1, \gamma_2, \alpha, \beta$ are sensible and γ_1 misses γ_2. Then the (α, β)-loops γ_1', γ_2' constructed by the surgery method above satisfy $i(\gamma_1', \gamma_2') \leq 4$. Furthermore, if γ_1 misses α, or β, then γ_1' misses α, or β, respectively.* □

Lemma 2.4. *Let α' be a component of a multicurve α on S and let β be a curve on S. Suppose α', β fill S. Then there exists a $(4,0)$-quasigeodesic $\alpha' = \gamma_0, \gamma_1, ..., \gamma_n = \beta$ with γ_i a (α, β)-loop for every $0 < i < n$.*

Proof. Start with a geodesic $\gamma_0, ..., \gamma_m$ of curves from α' to β, such that this collection of curves is sensible. Using the surgery process on each γ_i with $1 \leq i \leq n-1$, we obtain a sequence $\gamma_1', ..., \gamma_{m-1}'$ of (α, β)-loops. We have $i(\gamma_i', \gamma_{i+1}') \leq 4$ for each i by Lemma 2.3, therefore $d_S(\gamma_i', \gamma_{i+1}') \leq 4$ and $d_S(\gamma_i', \gamma_j') \leq 4|i-j|$ for each i, j. If for some $i > j$ we have $i - j > d_S(\gamma_i', \gamma_j')$, then we connect γ_i' and γ_j' with a geodesic and surger each vertex of it using α, β again. Repeating this process, we obtain the required quasigeodesic of (α, β)-loops. □

We remark that if $S \neq S_{1.2}$ then in Lemma 2.4 we can take a $(3,0)$-quasigeodesic, and for all but finitely many surfaces we can take a $(2,0)$-quasigeodesic.

Lemma 2.5. *Let Y be a subsurface of S and suppose ∂Y and β fill S. Let γ be a $(\partial Y, \beta)$-loop that cuts ∂Y. Then $d_Y(\gamma, \beta) \leq 2$ if Y is non-annular and $d_Y(\gamma, \beta) \leq 5$ otherwise.*

Proof. If Y is non-annular then any pair of arcs in the projection will intersect at most twice by Definition 2.1, so one can consider a closed regular neighbourhood of the arcs to prove the required bound on distance.

If Y is annular then suppose for contradiction that $d_Y(\gamma, \beta) \geq 6$. Then there exist arcs $\delta^* \in \pi_Y(\gamma)$ and $\epsilon^* \in \pi_Y(\beta)$ with $|\delta^* \cap \epsilon^*| \geq 5$. Following a claim from [11, Section 10], if we isotope the triangles cobounded by $\partial Y, \beta, \gamma$ into Y (this retains minimal position), we have that $|\delta^* \cap \epsilon^* \cap Y'| \geq 3$, where Y' is the homeomorphic lift of Y. Therefore there exists an arc of $\beta - \partial Y$ which intersects γ at least two times with the same sign, contradicting γ a $(\partial Y, \beta)$-loop. □

3. The proof

Let γ be a curve and P be a set of curves. We say γ is *ϵ-close* to P if for some curve β of P we have $d_S(\gamma, \beta) \leq \epsilon$. Throughout this section, δ is a constant such that $\mathcal{C}(S)$ is δ-hyperbolic, see Theorem 1.1.

Lemma 3.1. *There exists $D = D(\delta)$ such that for any subsurface Y, component $\alpha \subset \partial Y$, and geodesic $\alpha = \gamma_0, \gamma_1, ..., \gamma_n = \beta$ with $n \geq 3$, we have $d_Y(\gamma_i, \beta) \leq D$ whenever $i \geq 2$.*

Proof. Use Lemma 2.4 to construct a (4,0)-quasigeodesic Q of $(\partial Y, \beta)$-loops from α to β. For each i, we have γ_i is D'-close to Q where $D' = D'(\delta)$. For an explicit D', we can take $D' = D'' + 2$, where D'' is the largest integer with $D'' \leq \delta \lceil \log_2(26D'') \rceil$. See for example [3, Chapter III.H]. Using Lemma 1.2, we can take $D = 2D' + B$, where B is the bound provided in Lemma 2.5. □

Theorem 3.2. *Given a surface S there exists $M = M(\delta)$ such that whenever Y is a subsurface and $g = (\gamma_i)$ is a geodesic such that γ_i cuts Y for all i, then $d_Y(g) \leq M$.*

Proof. Take $M = 4\delta + 2D + 4$, where D is defined as in Lemma 3.1. Fix $i < j$. We shall show that $d_Y(\gamma_i, \gamma_j) \leq M$. Fix α a component of ∂Y. Let $I = N_{\delta+1}(\alpha) \cap g$. There exists $g' = (\gamma_{i'}, ..., \gamma_{j'})$ a geodesic of length at most $2\delta + 2$ such that $I \subset g' \subset g$.

Let P be a geodesic from α to γ_i and Q be a geodesic from β to γ_j. Let $i'' = \max\{i, i' - 1\}$ and $j'' = \min\{j, j' + 1\}$. Since geodesic triangles are δ-slim, we have either $\gamma_{i''}$ is δ-close to P and $\gamma_{j''}$ is δ-close to Q, or, there exists adjacent vertices of $g - g'$ with one δ-close to P and the other δ-close to Q. By lemmas 1.2 and 3.1 we have that $d_Y(\gamma_i, \gamma_j) \leq D + \delta + (2\delta + 4) + \delta + D = M$. □

We remark that M need not be optimal for each surface. For example, for S_2 it may be better to consider Leasure's cycles, which give $(2, 0)$-quasigeodesics, whereas a priori we are taking $(3, 0)$-quasigeodesics in Lemma 2.4. Similar surgery arguments may produce better results for other surfaces. Also, for all but finitely many surfaces, in Lemma 2.4 we can take a $(2, 0)$ quasigeodesic; this improves on the constant D.

4. Generalization to markings

We thank Brian Bowditch for suggesting this generalization and set-up. Defining the markings that we wish to discuss has similarities with [10, Section 6].

Given a multicurve α and a curve β such that α, β fill S, let B be a maximal collection of pairwise non-isotopic arcs of $\beta - \alpha$ in $S - \alpha$. We let $\Gamma_\alpha(\beta)$ be the graph embedded in S by taking the union $\alpha \cup B$. This may not be well-defined but there is bounded intersection between two such graphs, in terms of S, between any pair of choices of B.

Since α, β fill S, it follows that there are no essential simple closed curves on S that are disjoint from $\Gamma_\alpha(\beta)$, i.e. $\Gamma_\alpha(\beta)$ fills S. Furthermore, by an Euler characteristic argument, the number of edges of $\Gamma_\alpha(\beta)$ can be bounded in terms of the surface S. Let k_1 be this bound.

We write $\mathcal{M}_k(S)$ to denote the set of (isotopy classes of) embedded graphs that fill S with at most k edges. Let $\mathcal{M}_{k,l}(S)$ be the graph with vertex set $\mathcal{M}_k(S)$ with two vertices G_1, G_2 adjacent if $i(G_1, G_2) \leq l$. Here, $i(G_1, G_2) = \min |\Gamma_1 \cap \Gamma_2|$ where the minimum is taken over representatives Γ_i of the isotopy classes G_i, where $i = 1, 2$.

Let k_2 be a bound for the number of edges of any *clean complete marking* on S regarded as a graph on S, see [9] for definitions. The graph of clean complete markings on S is connected. Write $l_1 = \max_M i(M, M')$, where the maximum is taken for all clean complete markings of M, where M' differs from M by an *elementary move.* Let $l_2 = \max_G \min_M i(G, M)$, where the minimum is taken over graphs with at most $k = \max(k_1, k_2)$ edges that fill S and the maximum is taken over clean complete markings of S.

We then have $\mathcal{M}_{k,l}(S)$ connected, where $l = \max(l_1, l_2)$. Endow the graph $\mathcal{M}_{k,l}(S)$ with a metric where each edge has unit length, and distance is given by shortest paths. Vertex stabilizers are uniformally bounded, by the Alexander method. The *mapping class group* $\mathcal{MCG}(S)$ acts on $\mathcal{M}_{k,l}(S)$, and thus by the Milnor-Švarc Lemma [3, Proposition I.8.19] the mapping class group is quasi-isometric to the *marking graph* $\mathcal{M}_{k,l}(S)$.

Theorem 4.1. *Suppose a multicurve* α *and a geodesic* $g = (\gamma_i)$ *satisfy* γ_i, α *fill* S *for each* i*. Then* $\mathrm{diam}_{\mathcal{M}_{k,l}(S)}(\Gamma_\alpha(g)) \leq M$*, where* M *depends on* S*.*

Proof. We sketch a proof for brevity, since most of the proof is a generalization of earlier lemmas. Firstly, if Γ_1 intersects Γ_2 boundedly many times, then there are only finitely many possibilities for Γ_2 in terms of Γ_1. There are only finitely many possibilities for Γ_1 modulo homeomorphism. Thus, if intersection between markings is bounded then their distance is bounded.

Secondly, one bounds $i(\Gamma_\alpha(\gamma_1), \Gamma_\alpha(\gamma_2))$ when γ_1 and γ_2 are disjoint, in terms of S. This generalizes Lemma 1.2. Then one bounds $i(\Gamma_\alpha(\beta), \Gamma_\alpha(\gamma))$ when γ is a (α, β)-loop, in terms of S. This generalizes Lemma 2.5. Using these lemmas, one can generalize Lemma 3.1 then finish the argument analogously to Theorem 3.2. □

Acknowledgements. The author would like to thank Saul Schleimer for thorough comments on the paper. We thank Brian Bowditch, Saul Schleimer and Robert Tang for interesting conversations.

References

[1] T. Aougab, *Uniform hyperbolicity of the graphs of curves.* http://arxiv.org/abs/1212.3160.

[2] B.H. Bowditch, *Uniform hyperbolicity of the curve graphs.* http://homepages.warwick.ac.uk/~masgak/papers/uniformhyp.pdf.

[3] Martin R. Bridson and André Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 319, Springer-Verlag, Berlin, 1999. MR1744486 (2000k:53038)

[4] M.T. Clay, K. Rafi, and S. Schleimer, *Uniform hyperbolicity of the curve graph via surgery sequences.* in preparation.

[5] Benson Farb and Dan Margalit, *A primer on mapping class groups*, Princeton Mathematical Series, vol. 49, Princeton University Press, Princeton, NJ, 2012. MR2850125 (2012h:57032)

[6] W. J. Harvey, *Boundary structure of the modular group*, Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978), 1981, pp. 245–251. MR624817 (83d:32022)

[7] S. Hensel, P. Przytycki, and R. C. H. Webb, *Slim unicorns and uniform hyperbolicity for arc graphs and curve graphs.* http://arxiv.org/abs/1301.5577.

[8] Jason Paige Leasure, *Geodesics in the complex of curves of a surface*, ProQuest LLC, Ann Arbor, MI, 2002. Thesis (Ph.D.)–The University of Texas at Austin. MR2705485

[9] H. A. Masur and Y. N. Minsky, *Geometry of the complex of curves. II. Hierarchical structure*, Geom. Funct. Anal. **10** (2000), no. 4, 902–974. MR1791145 (2001k:57020)

[10] Howard Masur, Lee Mosher, and Saul Schleimer, *On train-track splitting sequences*, Duke Math. J. **161** (2012), no. 9, 1613–1656. MR2942790

[11] Howard Masur and Saul Schleimer, *The geometry of the disk complex*, J. Amer. Math. Soc. **26** (2013), no. 1, 1–62. MR2983005

[12] Howard A. Masur and Yair N. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. **138** (1999), no. 1, 103–149. MR1714338 (2000i:57027)

Mathematics Institute, University of Warwick, Coventry, CV4 7AL, United Kingdom.
E-mail address: R.C.H.Webb@warwick.ac.uk